

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-092491

(43)Date of publication of application : 06.04.2001

(51)Int.Cl. G10L 15/20  
G10L 21/02  
G10L 15/02  
H04R 3/02  
// G10L101:027

(21)Application number : 2000-265121

(71)Applicant : TRW INC

(22)Date of filing : 01.09.2000

(72)Inventor : LAMBERT RUSSELL H  
EDMONDS KARINA L  
HSU SHI-PING

(30)Priority

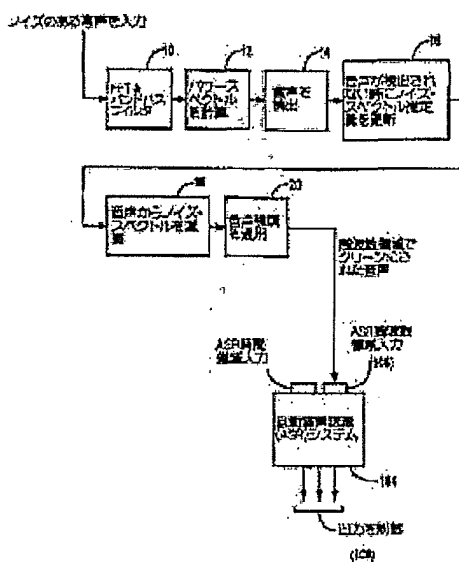
Priority number : 1999 388266 Priority date : 01.09.1999 Priority country : US

## (54) SYSTEM AND METHOD FOR REDUCING NOISE BY USING SINGLE MICROPHONE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a noise reduction system used with a single microphone channel.

SOLUTION: This noise reduction system comprises a band-pass filter 10 for eliminating known noise frequencies from a speech signal, a noise detector 14, a noise estimating unit 16 for updating a noise estimated amount only when no speech is detected, a spectrum subtraction circuit 18 for subtracting the noise estimated amount from the speech and noise signal spectrum, a speech emphasis circuit 20 for further emphasizing speech signal components related to some residual noise, and also a fast Fourier transform circuit 10. The noise-reduced signal in the frequency domain obtained as a result is directly inputted to an automatic speech recognition system 104 or re-transformed into the time domain for the use in a speech communication system.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-92491

(P2001-92491A)

(43) 公開日 平成13年4月6日 (2001.4.6)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
G 1 0 L 15/20		H 0 4 R 3/02	
21/02		G 1 0 L 101:027	
15/02		3/02	3 0 1 D
H 0 4 R 3/02			3 0 1 G
// G 1 0 L 101:027		7/08	A

審査請求 有 請求項の数 7 O L (全 9 頁) 最終頁に続く

(21) 出願番号 特願2000-265121(P2000-265121)

(22) 出願日 平成12年9月1日 (2000.9.1)

(31) 優先権主張番号 09/388266

(32) 優先日 平成11年9月1日 (1999.9.1)

(33) 優先権主張国 米国 (U S)

(71) 出願人 591169755

ティーアールダブリュー・インコーポレー  
テッド

TRW INCORPORATED

アメリカ合衆国オハイオ州44124, リンド  
ハースト, リッチモンド・ロード 1900

(72) 発明者 ラッセル・エイチ・ランバート

アメリカ合衆国カリフォルニア州92708,  
ファウンテン・ヴァリー, コロンバス・コ  
ート 9501

(74) 代理人 100089705

弁理士 社本 一夫 (外5名)

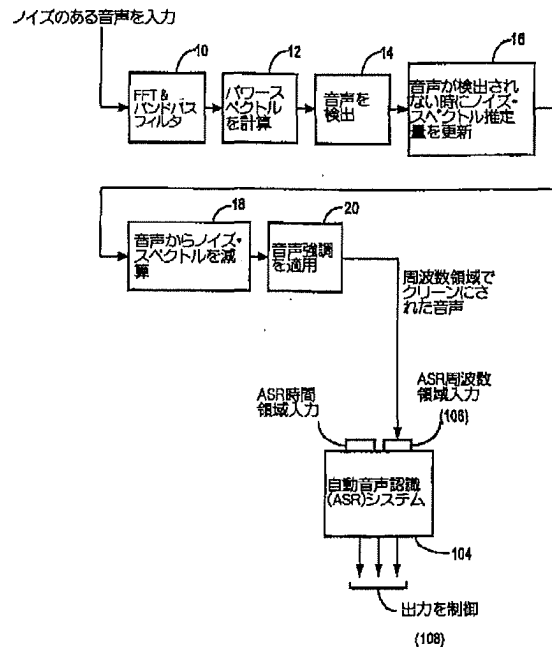
最終頁に続く

(54) 【発明の名称】 単一マイクロフォンを使用するノイズ減少のためのシステムおよび方法

(57) 【要約】

【課題】 単一マイクロフォンチャンネルと共に使用する  
ためのノイズ減少システムを提供する。

【解決手段】 音声信号から既知のノイズ周波数を除去す  
るバンドパスフィルタ10、ノイズディテクタ14、音  
声を検出されなかった時のみノイズ推定量を更新するノ  
イズ推定器16、音声とノイズ信号スペクトルからノイ  
ズ推定量を引くスペクトル減算回路18、いくらかの残  
存ノイズに関連した音声信号成分をさらに強調する音声  
強調回路20と共に高速フーリエ変換回路10を含む。  
結果として得られた、周波数領域のノイズ減少信号は直  
接、自動音声認識システム104に入力されるか、音声  
通信システム内で使用するために時間領域へ再変換され  
る。



## 【特許請求の範囲】

【請求項1】 入力マイクロフォン・データ・ブロックを周波数領域に変換する高速フーリエ変換（FFT）回路と、

ノイズが存在していると知られている、選択された周波数帯域を取り除く帯域フィルタと、  
マイクロフォン・データ中の音声信号の存在を検出する音声ディテクタと、

音声信号が検出されないデータ・ブロックのためのみに更新されるノイズ・スペクトル推定器と、

ノイズと音声信号成分を含むマイクロフォン信号から、前記推定されたノイズ・スペクトルを減算するためのスペクトル減算回路と、

周波数領域においてノイズ減少音声信号を提供するため、前記スペクトル減算回路の動作後の何らかの残存ノイズに関し、前記音声信号成分を強調するための音声強調回路と、を含む、ノイズ環境における単一マイクロフォンのためのノイズ減少システム。

【請求項2】 前記周波数領域においてノイズ減少された音声信号から時間領域データを再構成する手段をさらに含み、該手段は、前記周波数領域から時間領域へデータ・ブロックを変換し直すための逆高速フーリエ変換回路を含み、それによって前記ノイズ減少された音声信号は、音声通信システムにおいてさらに明瞭になる、請求項1に記載されたノイズ減少システム。

【請求項3】 前記周波数領域において前記ノイズ減少された音声信号を受取るために結合された自動音声認識（ASR）システムを含み、それによって前記ASRシステムは、選択された制御信号を生じるのにより信頼性をもって動作する、請求項1に記載されたノイズ減少システム。

【請求項4】 入力信号が、各「A」サンプルのブロックにおけるノイズ減少システムに与えられ、各サイズ「2A」サンプルのデータ・ブロックが、前記FFT回路へ与えられ、

データ・ブロックを形成するために、「A」サンプルの入力信号ブロックを2つ1組に結合するための手段をさらに含み、

入力信号ブロックを結合するための前記手段は、現在の入力信号ブロックが、現データ・ブロックの後半に、それから次のデータ・ブロックの前半に置かれるように、各入力信号ブロックを2回使用し、

各データ・ブロックに対して三角形の重み付けウィンドウを適用する方法をさらに含み、

時間領域データを再構成する前記手段は、各再構成データ・ブロックの前半と、前のデータ・ブロックの処理から保存された再構成データ・ブロックの後半とを結合するための手段を含み、均一なエンヴェロープを有する時間領域サンプルが再構成されそしてブロック処理の望ましくないアーティファクトが最小化される、請求項2に

記載されたノイズ減少システム。

【請求項5】 単一マイクロフォンからの入力データ・ブロックを、時間領域の表示から周波数領域の表示へ変換するステップと、

ノイズ源と知られた効果を最小にするために、選択された周波数帯域をフィルタリングするステップと、

各ブロックのデータ信号における音声の存在を検出するステップと、

音声を検出されない時、ノイズ・スペクトル推定量を更新することによりノイズを推定するステップと、

10 入力音声およびノイズ信号から、前記ノイズ・スペクトル推定量を減算するステップと、

周波数領域音声信号に減算されたノイズ要素を提供するために、前記減算ステップの結果をN番目のパワーまであげることによってノイズ信号成分に関して音声信号成分を強調するステップと、ここでNは1より大きな正の量である、を含む、ノイズ環境において単一マイクロフォンによって生じた信号内のノイズを減少させる方法。

【請求項6】 前記周波数領域からのデータ・ブロックを時間領域に変換しステップを含み、周波数領域の前記ノイズ減少音声信号からの時間領域信号を再構成ステップをさらに含み、それによって音声通信システムにおいて前記ノイズ減少音声信号がより明瞭になる、請求項5に記載された方法。

【請求項7】 各「A」サンプル・ブロックにおいて入力信号を前記ノイズ減少システムに与えるステップと、各サイズ「2A」サンプルのデータ・ブロックを前記FFT回路に与えるステップと、

30 「A」サンプルの入力信号ブロックを2つ1組に結合してデータ・ブロックを形成するステップと、前記結合ステップは、現在の入力信号ブロックが現データ・ブロックの後半にそれから次のデータ・ブロックの前半に置かれるように、各入力信号ブロックを2回使用するステップを含み、

三角形の重み付けウィンドウを各データ・ブロックに適用するステップと、をさらに含み、

前記再構成ステップにおいて、前のデータ・ブロックの処理から保存された再構成データ・ブロックの後半に各再構成データ・ブロックの前半を結合するステップをさらに含み、均一なエンヴェロープを有する時間領域サンプルは再構成され、ブロック処理の望ましくないアーティファクトは最小化される、請求項6に記載された方法。

## 【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、音響上騒がしく反響する環境において、音声データを音響信号から電気的信号に、確実に変換するための技術に一般的に関する。

【0002】

50 【従来の技術】 ダイヤルおよび他の機能のための自動音

声認識 (A S R) を使用し、自動車からの「ハンドフリー・セル」携帯無線電話通信に対する需要が高まっている。しかし、車両内で通信は、車内および車外からの背景ノイズによって、困難およびストレスの多いものとなっている。自動車内での反響は高いノイズレベルと結合し、自動車内におけるマイクロフォンにより受取られた音声信号を大きく減衰させる。マイクロフォンは、原音声信号だけでなく、自動車内部の壁、窓および物体からの多重エコーから生じた、歪んだそして遅延した音声信号の複製をまた受信する。これらの複製信号は一般に、異なった経路を通して、マイクロフォンに到着する。それゆえ、その環境に対して、「多重経路」という語がしばしば付けられる。そのような環境の中では、音声信号の質が極端に低下し、何らかの関連した A S R システムの精度 (正確度) は、もはや動作しない点にまで低下してしまう。例として、静かな環境内において 96 % の高さの認識精度のある A S R システムが、走行している自動車の中ではその精度は 50 % 以下まで大きく落下する。

【0003】ノイズおよび反響に影響される別の関連技術は、通信帯域における減少を達成するためおよび他の理由のために、音声信号をデジタル的に符号化する、音声圧縮である。ノイズの存在で、音声圧縮はますます難しく、信頼性のないものになる。

【0004】

【発明が解決しようとする課題】音響分野においてアクティブ・ノイズ・キャンセレーションを果たす多くの先行技術システムがある。アクティブ・ノイズ減少方法は、望ましくないノイズ信号をキャンセルするために、ノイズ源近くの、1 以上の変換器を通して、時として「アンチノイズ」として参照される反対信号を生じることによって音響ノイズ信号をキャンセルする。この技術では、しばしば、スピーカの付近のいくつかの他の位置にノイズを作出し、特に多重経路効果の存在のため、多重の未知のノイズ源をキャンセルする実践的な解決とはいえない。

【0005】従って、走行自動車の内部のような反響する環境内で、ノイズ効果を減少させる重要な要求がまだある。以下に議論するように、本発明はこの要求に取り組んでいる。

【0006】

【課題を解決するための手段】本発明は、騒音環境において、単一マイクロフォンから得られる音声信号内のノイズを減少するシステムおよび方法である。本発明は、如何なる所与のアプリケーション、ノイズ環境もしくは、自動音声認識 (A S R) システムに対し、多重パラメータを最適に調整できる、一般的なノイズ減少構造に関するものである。簡単に、一般的な言葉では、本発明のシステムは、入力されたマイクロフォン・データ・ブロックを周波数領域表示に変換するための高速フーリエ

変換 (F F T) 回路; ノイズが中に存在していることが知られている、選択された周波数バンドを排除するためのバンドパス・フィルタ; マイクロフォン・データにおける音声信号の存在を検出する音声ディテクタ; 音声信号が検出されないデータ・ブロックのためにのみ更新されるノイズ・スペクトル推定器、ノイズと音声信号成分を含んでいるマイクロフォン信号から推定されたノイズ・スペクトルを減算するためのスペクトル減算回路; 周波数領域においてノイズ減少された音声信号を提供するために、スペクトル減算回路の動作後の何らかの残存ノイズに関して、音声信号成分を強調するための音声強調回路; を含む。

【0007】システムはさらに、周波数領域内でノイズ減少された音声信号から、時間領域データを再構成し、周波数領域からのデータ・ブロックを時間領域に変換し直す逆高速フーリエ変換回路を含み、それによって音声通信システムにおけるノイズ減少された音声信号をより理解しやすくなる。さもなければ、システムは、周波数領域のノイズ減少音声信号を受け取るように結合されている自動音声認識 (A S R) システムをさらに含むことができ、それによって、A S R システムは選択された制御信号を生じるために、より正確に動作する。

【0008】好ましくは音声強調回路は、パワー  $N$  によって周波数領域の信号を高める、ここで  $N$  は 1 より大きな正の量である。開示された発明において、入力信号は “A” サンプルの個々のブロックにおいてノイズ減少システムに与えられ、サイズ “2 A” サンプルの個々のデータ・ブロックは F F T 回路に与えられる。システムは、“A” サンプルの入力信号ブロックを 2 つ 1 組に結合してデータ・ブロックを形成するための手段をさらに含む。さらに、入力信号ブロックを結合するための手段は、現在の入力信号ブロックが現データ・ブロックの後半にそして次のデータ・ブロックの前半に置かれるように、各入力信号ブロックを二回使用する。システムは、各データ・ブロックに対して、三角の重み付けウインドウを適用するための手段をさらに含み; 時間領域データを再構成するための手段は、以前のデータ・ブロックを処理することから保存された、再構成データ・ブロックの後半と各再構成データ・ブロックの前半を結合するための手段を含み、一様なエンヴェロープを有する時間領域サンプルは再構成され、ブロック処理動作の望ましくないアーティファクトは最小限になる。

【0009】本発明の別の側面に従うと、本システムは、いつノイズ減少を利用するのが望ましいのかという指示を与えるノイズモニタ; ノイズモニタにおいて検出されたノイズレベルが比較的高いとして検出された時に、ノイズ減少信号を選択し、および検出されたノイズ・レベルが比較的低いときに、ノイズ信号を有する原音声を選択する手段を、さらに含む。

【0010】本発明は、ノイズ環境において、単一マイ

クロフオンによって受信された信号中のノイズを減少させるための方法に関して限定されう。簡単に一般的な言葉で言うと、本方法は、単一マイクロフォンからの入力データ・ブロックを、時間領域表示から周波数領域表示に変換するステップを含み；ノイズ源と知られた効果を最小限にするため、選択された周波数帯域をフィルタ作用し；データ信号の各ブロック中の音声の存在を検出し；音声を検出されない時に、ノイズ・スペクトル推定量を更新することでノイズを推定し；入力音声およびノイズ信号からノイズ・スペクトル推定量を減算し；ノイズ信号成分に関して音声信号成分を強調し；減算するステップの結果をN番目、ここでNは1より大きな正の量である、のパワーまで上昇させ、減少したノイズ含有量を持つ周波数領域音声信号を提供する。

【0011】本方法は、周波数領域におけるノイズ減少された音声信号から時間領域データを再構成するステップをまた含み、周波数領域からのデータ・ブロックを時間領域へ戻し変換するステップを含み、それにより、音声通信システムにおいてノイズ減少音声信号は、より明瞭になる。さもなければ、本方法は周波数領域でノイズ減少された音声信号を、自動音声認識（ASR）システムへ伝送するステップを含み、それによって、ASRシステムは、選択された制御信号を生じるために信頼性をもって動作する。

【0012】好ましくは、音声信号成分を強調する方法ステップはパワーN、ここでNは1より大きな正の量である、により周波数領域において信号を上昇するステップを含む。

【0013】さらに特定すると、本方法は、各“A”サンプルのブロックにおいて入力信号をノイズ減少システムに与え；サイズ“2A”サンプルのデータ・ブロックをFFT回路に与え；データ・ブロックを形成するために“A”サンプルの入力信号ブロックを2つ1組に結合し、その結合ステップは、現在の入力ブロックが現在のデータ・ブロックの後半におよび次のデータ・ブロックの前半に置かれるように、各入力信号ブロックを二回使用する過程を含み；各データ・ブロックに対して、三角の重み付けウィンドウを適用し；再構成ステップにおいて、以前のデータ・ブロックを処理することから保存された再構成データ・ブロックの後半と、各再構成データ・ブロックの前半を結合するステップをさらに含む。均一エンヴェロープを有する時間領域サンプルは再構成され、そしてブロック処理動作の望ましくないアーティファクトはこの方法の使用により最小化される。

【0014】本方法は、いつノイズ減少を利用するのが望ましいかという指示を与えるために、ノイズ・モニタにより連続的にノイズ・レベルを監視し；ノイズ・モニタによって検出されたノイズ・レベルが比較的高く検出された時、ノイズ減少信号を選択し；検出されたノイズレベルが比較的低い時、原音声およびノイズ信号を選択

する；各ステップをさらに含みうる。

【0015】前述の要約から、本発明はノイズ減少技術における重要な発展を表わしている、と理解されるであろう。上記において要約された特色を組み合わせると、ノイズが著しく減少した音声信号となり、その結果、音声通信システムで信号が使用される時にさらに明瞭な音声となり、信号がASRおよびそれに関連したシステムのために動作するよう使用される時に、もっと信頼性のあるASRシステム動作となる。発明の他の側面および利点は、以下に続くさらに詳細な説明および添付した図と共に明らかになるであろう。

【0016】

【発明の実施の形態】図に示されているように、本発明は、動いている自動車の内部のような、ノイズがあり反響する環境における音声の検出中に、ノイズ効果を大いに減少するための技術に関係している。自動車内の移動電話からの音声伝送の音質は殆どずっとよくないことが長い間知られてきた。車両の内外からのノイズにより、信号対雑音比が比較的低くなり、車両内での音響の残響が音声信号を低下させる。自動音声認識（ASR）および音声圧縮のために利用できる技術は、いくらよくみても低下しており、自動車の環境の中では全然作用し得ない。

【0017】本発明に従うと、図1に示すように、ノイズのスペクトル減算を含む処理ステップの組合わせは、ノイズレベルの重要な減少を達成するために実行される。ノイズある音声信号はデジタルサンプルに変換され、ブロック10の中で示したように、高速フーリエ変換（FFT）回路内で処理するために、同時にサンプル・ブロックに入力される。ブロック10にも示すように、高速フーリエ変換による周波数領域への変換により、信号はまず帯域フィルタされる。それから、ブロック12の中で示すように、FFT関数の絶対値として強度スペクトルが計算される。次に、まだ周波数領域にある各ブロックのデータは、ブロック14の中で示すように、音声の存在、不存在を検出するために分析される。発明の本質的な部分は、ノイズ・スペクトル推定量のスペクトル減算によってノイズを減少することにある。理想的には、この推定は、音声が存在していない時に得られたデータに基づかれるべきである。ブロック16に示すように、もし音声が存在していれば、ノイズ・スペクトル推定量は更新されず、音声が存在していないならばノイズ推定量は更新される。

【0018】ブロック18に示されているように、ノイズ・スペクトル推定量は、まだ周波数領域にある、ノイズのある音声信号スペクトルから減算される。そしてブロック20に示したように、音声信号（ノイズのスペクトル減算の後に得られた）をn番目のパワー、ここでnはもっとも望ましい結果を提供するように最適化されている、まで上昇させることによって、音声をあらゆる残

存ノイズにわたりさらに強調する。最後に、ブロック図22で示すように、周波数領域のデータ・ブロックは、逆FFT回路によって逆変換させられ、時間領域で「クリーン」な音声を出力する。

【0019】図2は、図1に描かれた関数をさらに詳細＊

パラメータ名	記述	範囲	単位
A	ブロック・サイズ (FFTサイズは2A)	正の実整数(通常2の パワー)	サンプル
B	入力低カットオフ・ ポイント	0パラメータC	周波数(Hz)
C	入力高カットオフ・ ポイント	パラメータBサンプル・ レート/2	周波数(Hz)
D	スペクトル圧縮ファクタ	正の実数(1より大きい)	単位なし
E	音声位置低制限	0パラメータF	周波数(Hz)
F	音声位置高制限	パラメータEサンプル・ レート/2	周波数(Hz)
G	現在の平均エネルギー 更新パラメータ	正の実数(0と1の間)	単位なし
H	音声検出閾値パラメータ	正の実数	単位なし
I	走行平均ノイズスペクトル 更新パラメータ	正の実数(0と1の間)	単位なし
J	音声エンハンスメント・ パラメータ	正の実数(1より大きい)	単位なし

【0021】図2に示す関数は、いくつかの望ましいハードウェアもしくはソフトウェアの形態のなかで履行されうる。試験的な形態として、ノイズ・キャンセレーション・システムは、リアルタイムでパソコン上において動くマイクロソフト・ビジュアルC++コンパイラにおけるコードを持ったソフトウェアとして実現された。入力音声信号はサンプリ化され、各Aサンプルのブロックにおける入力である。FFT処理動作のための計算ブロックは、それぞれ2Aデータサンプルを含むように形成される。このため、FFTポイント・サイズは2Aになる。例えば、Aは128サンプルであり、2Aは256サンプルであり得る。

【0022】図2の長方形40は、データ・ブロックの入力を示している。長方形42は、2Aサンプルの各データ計算ブロックは、Aサイズブロックのストリームからオーバーラッピング方法において形成されることを示している。もっと明確に言うと、Aサイズ・ブロックの到来ストリームは、ブロック(a)、ブロック(b)、ブロック(c)、ブロック(d)等として表され、このとき第一のデータ計算ブロックは、ブロック(a)とブロック(b)から共に形成され、次のデータ計算ブロックはブロック(b)とブロック(c)から共に形成され、次はブロック(c)とブロック(d)から共に、等々と形成される。この方法でブロックをオーバーラップするのは、連続的にデータ・ブロックを処理することによって導入され得るサウンド・アーティファクトを最小限にするためである。さらに、各データ計算ブロックは、長方形44の中で示すように、データ計算ブロックの中心に置かれた二等辺三角形のプロファイルを有する三角形重みづけ関数によって、「ウィンドウ処理(windowing)」させられる。この結果、最大重みが、データ計算ブロックの中心にある一つもしくは複数のサンプルに適用され、より小さな重みが段々にブロッ

＊に描いている。図2で言及されている一般的パラメータの組は、以下の表1に規定されている。

【0020】

【表1】

クの先端および後端へ向かうサンプルへ適用される。データ計算ブロックは、Aサイズ・ブロックのオーバーラップ処理からデータを得るので、これらの三角ウィンドウはまたオーバーラップする。さらに、信号が後に周波数領域に変換されて時間領域にもどされるとき、各隣接する対のオーバーラップするデータ計算ブロックのからの寄与部分が結合して、比較的均一なサイズのエンヴェロープをもつサンプルの組を発生させる。

【0023】各連続したデータ・ブロックが形成されそしてウィンドウ化されると、長方形46に示したように、FFT処理動作へ導かれ、長方形48に示されるように、パラメータBとCによって規定された限界の間で帯域フィルタリングするように指向される。このフィルタリング・ステップによって、300Hz以下のような非常に低い周波数および3850Hz以上のような非常に高い周波数のノイズを除く。次に、長方形50で示すように、強度スペクトルSが計算されそしてパラメータDを使用する圧縮領域に置かれる。 $S_{compressed} = S^{1/D}$ 。

【0024】長方形52で示されたように、現在のデータ・ブロックの音声エネルギーは、パラメータEおよびFによって与えられた周波数範囲、音声をもっとも支配的になるうる、400Hzから800Hzような範囲で、エネルギーを加算することにより計算される。この範囲の平均音声エネルギーは長方形54に示したように、実行平均推定量に維持され、次の計算式を使用する：

【0025】

【数1】  $SpeechEnergy_{avg} = (1-G) * SpeechEnergy_{avg} + G * SpeechEnergy_{current}$

決定ブロック56において、現音声エネルギーは、連続的に適合している音声検出閾値を生じる、平均音声エネルギー $E_{avg}$ をH倍したものと比較される。現音声エネルギーがその $H * E_{avg}$ より大きいならば、経路58によって示

されているように、ノイズ・スペクトルは更新されない。そうでなければ、ノイズ・スペクトルは長方形60で示されているように、次の式を使用するパラメータIを使用して更新される：

【0026】

【数2】

$$\text{Spectrum}_{\text{avg}} = (1-I) * \text{Spectrum}_{\text{avg}} + I * \text{Spectrum}_{\text{current}}$$

音声スペクトルは、長方形62に示すように、現在のスペクトルとノイズ・スペクトル推定量との差として計算される。最後に、重要な音声エンハンスメント・ステップ64があり、残存ノイズ構成要素を含む音声スペクトルはパワーJ、Jは1より大きくなるよう選ばれている、まで高められる。1より大きなパワーまで信号を高めることで、さらに、音声構成要素をノイズ構成要素から区別することができる。

【0027】パラメータ最適化の例として、パラメータJの様々な値の結果が観測され（その間すべての他のパラメータは固定状態）、以下に表2で表される：

【0028】

【表2】

音声エンハンスメント・パラメータJ	ASRからの精度
1.5	80%
1.7	81.4%
1.85	84%
1.9	85.6%
1.95	81.4%
2.0	80.7%
2.2	76.4%
2.5	67.1%

【0029】自動音声認識の観点からすると、パラメータJの最適値は1.9と分かるであろう。音声信号がシステムの人間のユーザに伝えられるならば、それらは次に時間領域に変換されなければならない。時間領域波形の再構成も、ブロックごとを基準に実行される。逆FFT動作は、長方形66に示されているように、各データ・ブロックで実行される。その結果の、三角形にウィンドウ化されたデータ・サンプルは、再構成波形の均一データ・エンヴェロープを生じる方法で共に加算されなくてはならない。もっと特定すると、ブロック68に示すように、再構成されたデータ・ブロックの前半は前に変換されたデータ・ブロックの後半に加えられる。これらの2つの半分のブロックは、当初に三角ウィンドウ化するようにされたので、それらは均一の信号エンヴェロープを生じるような相補的な方法で、今度は結合する。長方形70に示すように、現ブロックの後半は、次のブロック繰返しのために保存される。現在および以前のブロックから結合されたAサンプルは、長方形72に示すように、出力される。

【0030】最良の動作のためには、標準的な“スター・サーチ”（star search）技術が、上述の方法の一つのパラメータを変えて、他のすべてのパラメータを固定しながら利用できる。理想的には、これは、

各音声タイプおよび異なるノイズ条件のために繰返されなければならない。もっとも重要なパラメータの一つは音声強調項Jである。これは、Jの各設定のための認識精度をテストしている間、1.5から2.5まで変化させられる。示された最適パラメータ値は、高速道路および車両ノイズの存在する場所において本発明を使用し、そして音声で結合されたデジット・データのためであった。

【0031】図3に示すように、グラフ80によって示された、ランダム・ノイズは、サイン波が周期的な自己相関関数を有するのに対し、自己相関関数82のなかに、特徴的な「スパイク」をもっている。音声84のセグメントは、周期的なサイン波である強い成分を持っている。そこで、86に示すように、音声は数ミリ秒にわたり強く相関する。対照的に、自己相関関数82内のスパイクによって示されるように、ノイズ80はゼロ遅延点のみで強く相関する。相関領域では、ノイズによるスパイクは、容易に消去されることができ、これが本発明の中で利用されているスペクトル減算方式の基礎となっている。

20 【0032】本発明のシステムは、窓を閉じてエアコンをつけそして窓を部分的に開けて高速道で走行中の車両における実践的条件下でテストされてきた。全方向、単方向の二つのタイプのマイクロフォンが考案された。予期に反せずに、単方向マイクロフォンは、すべての背景ノイズ・レベルに対して、かなりより高い認識精度を導いた。得られた最高の認識精度は、結合したデジタル音声データを使用中に、窓を上げそしてエアコンを点けて高速道路ドライブ中の場合の86%であった。

30 【0033】車内でのデータは、始めにデジタル・レコーダを使って集められ、マイクロフォンの配置は、信号対雑音比（SNR）が最大になるように選ばれる。全方向と単方向の両方のマイクロフォンにとって、最大の信号を生じる位置は、ドライバーの遮光版（すなわち、ソースの直前）の直ぐ上である。すべてのテストは、音声のポイント・ソースとして乗客を使って行われた。車の客席は対称的なので、ドライバ側の結果は、乗客側から得られた結果と等しくなることが予想できる。自動車の中でデジタル・レコーダに録音された音声は、44.1 kHzでサンプル化され、続けて8 kHzまでダウンサンプルした。ダウンサンプリング後のオーディオファイルの完全性を保証するために、そのファイルは自動音声認識（ASR）システムとともに試験された。44.1 kHzで記録されそして8 kHzまでのダウンサンプルされたファイルには、ASRにおけるデグラデーションは観測されなかった。

40 【0034】ASRシステムにおいて、認識精度はデジット誤差率により計算される。置換（S）、削除（D）、挿入（I）の数はテストされたデジット（N）のトータル数によって割られる。

50 【0035】

【数3】 $\text{Error} = (S+D+I)/N \times 100$

LernoutとHaupieによって設計されたソフトウェア・パッケージASR1500は、結合されたディジットに対して許容されそして比較的短い応答時間を持っているので、テストのために利用された。テストされた語彙は11のディジット、つまり1-9、0、オー（oh）について考慮される。認識過程で、同時調音ファクタを計算するために、結合されたディジットは選出される。テスト処理では、各ディジットはランダムに連続して接続されたディジットの会話の間約15回発音される。

【0036】背景ノイズの除去のあと、ディジットに対する認識精度は、非常に改善されている。窓を上げそしてエアコンを点けた状態では、単方向マイクロフォンの認識精度は47%から86%に改善され、全方向マイクロフォンの精度は16%から78%に改善された。窓を部分的に開けた状態では、単方向マイクロフォンの認識精度は46%から83%に改善され、全方向マイクロフォンの精度は10%未満から39%に改善された。

【0037】図4に示すように、背景ノイズ・レベル監視システム90は、発明の標準的なノイズ・キャンセレーション・システムに組入ることができ、この場合は背景ノイズの特定されたレベルが存在している時のみ動作するであろう。背景ノイズがない時、これは処理から音声デグラデーションを除くであろう。決定は“ハード”（オンもしくはオフ）のものである必要はない。高ノイズ条件における処理動作に変わる効果がシステムのユーザに気付かれないように連続的に変わる方法で、修正されたシステムは、処理された音声と処理されていない音声を適切に混ぜ合わせるであろう。例によると、本発明のこの実施形態において、決定ブロック92示されているように、モニタされたノイズレベルはさらに高い閾値に対して比較され、もしノイズがその閾値を超えるならば、長方形94に示されるように、システムは処理された（ノイズ減少された）音声を選択する。もしモニタされたノイズ・レベルが現在、より高い閾値より低いならば、決定ブロック96に示すように、より低い閾値と比較される。もしノイズがより低い閾値より低いならば、長方形98に示すように、元の未処理の音声を選ばれる。もし、モニタされたノイズがより高い閾値とより低い閾値との間ならば、長方形100に示すように、システムは元の音声およびノイズ減少された音声信号からの入力の

混合を選択する。

【0038】発明の他の実施形態として、ノイズ減少システムは自動音声認識システム（ASR）104（図5）に組込まれる。ノイズ減少システムは、図1に描いたものと同じであるが、最後の逆FFT処理を含んでいない。これは、時間領域波形に変換し直す時に作られるいくつかの音声アーティファクトを除去するであろう。アプリケーションがASRシステムの音声制御のみを必要とする場合には、時間領域波形を再構成する必要がない。逆フーリエ関数はノイズ・キャンセレーション・システムから除かれ、ノイズ・キャンセレーション・システムの出力は直接、ASRシステム104の周波数領域の入力106へ結合され、入力音声コマンドの検出に回答して、適切な出力制御信号108を生じる。

【0039】以上から明らかなように、本発明は、走行中の自動車のような、ノイズある環境に取付けられた単一マイクロフォンに対しノイズ減少の重要な発展を表明するものであると理解されるであろう。特に、本発明は、「クリーンにされた」すなわち人間の耳にとってより明瞭であるノイズ減少された音声信号を提供し、ASRシステムの信頼性を改善した。発明のシステムは、音声通信システム上で伝播するための時間領域出力もしくは、ASRシステムと直接結合するための周波数領域出力のいずれかを生成する。たくさんの実施形態が例示のために詳しく述べられているが、本発明の意図および範囲から離れることなく様々な修正を施し得ることも明らかである。従って、発明は添付された請求項以外を除いては制限されない。

【図面の簡単な説明】

【図1】本発明に従ったノイズ・キャンセレーション・システムのブロック図である。

【図2】発明のノイズ・キャンセレーション・システムのさらに詳細なブロック図である。

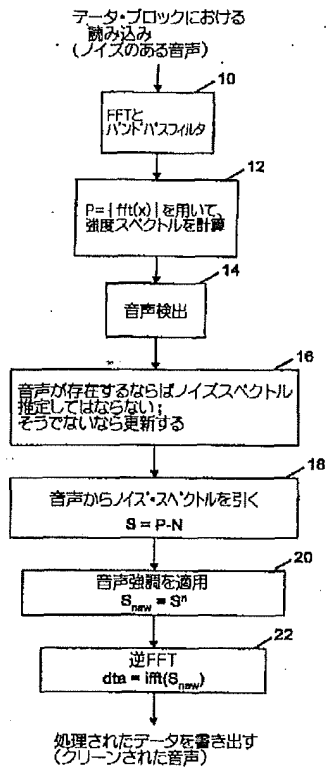
【図3】ノイズ信号自らの時間領域相関、即ち自己相関および音声信号の時間領域自己相関を示す、4つの関連したグラフの組である。

【図4】ノイズ・キャンセレーション・システムの動作を制御ために、ノイズディテクタを利用した発明の別の実施形態を表すブロック図である。

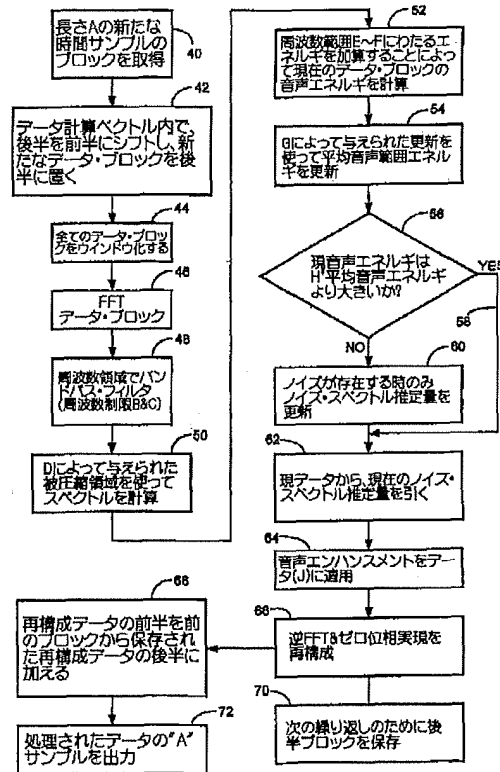
【図5】発明のノイズ・キャンセレーション・システムが、どのように現存の音声認識システム内（ASR）へ集積される得るかについて示すブロック図である。



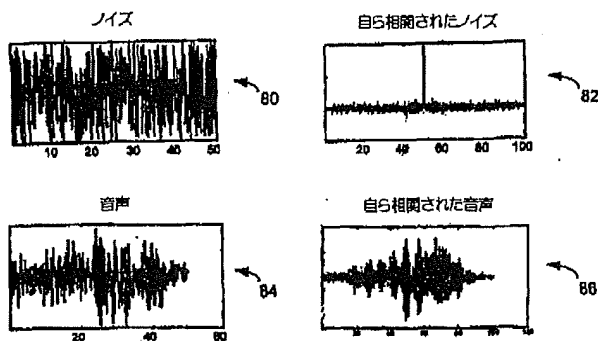
【図1】



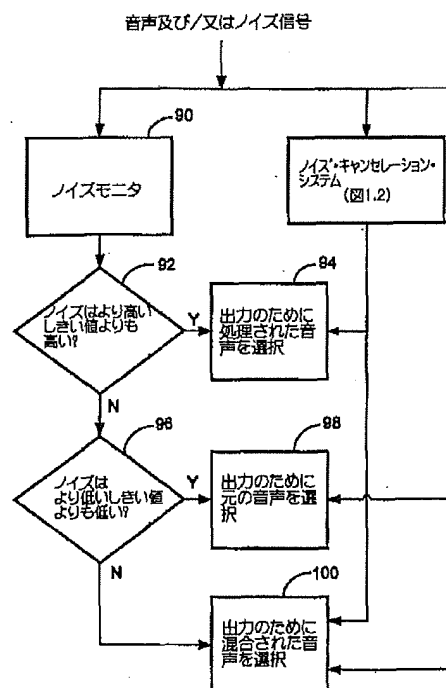
【図2】



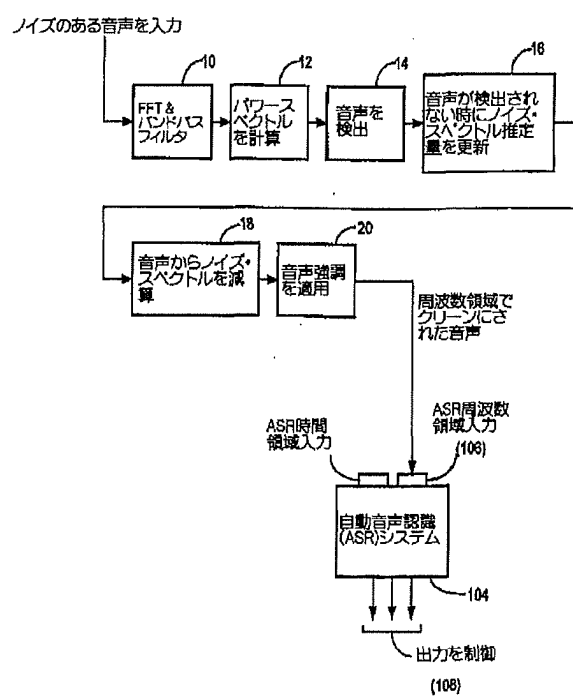
【図3】



【図4】



【図5】



フロントページの続き

(51) Int. Cl.<sup>7</sup>

識別記号

F I

7-マコード (参考)

G 1 0 L 9/00

F

(72)発明者 カリーナ・エル・エドモンズ  
アメリカ合衆国カリフォルニア州91106,  
パサデナ, コードヴァ・ストリート  
1160, ナンバー10

(72)発明者 シーピン・スー  
アメリカ合衆国カリフォルニア州91107,  
パサデナ, サウス・ボニータ・アベニュー  
461